HAL
open science

Semi-supervised spectral clustering with automatic propagation of pairwise constraints

Nicolas Voiron, Alexandre Benoit, Patrick Lambert, Andrei Filip, Bogdan Ionescu

To cite this version:

Nicolas Voiron, Alexandre Benoit, Patrick Lambert, Andrei Filip, Bogdan Ionescu. Semi-supervised spectral clustering with automatic propagation of pairwise constraints. International Workshop on Content-Based Multimedia Indexing (CBMI), Jun 2015, Prague, Czech Republic. pp.1-6, 10.1109/CBMI.2015.7153608 . hal-01229055

HAL Id: hal-01229055

https://hal.science/hal-01229055

Submitted on 16 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Semi-supervised Spectral Clustering with automatic propagation of pairwise constraints

Nicolas Voiron, Alexandre Benoit and Patrick Lambert
LISTIC, Université Savoie Mont Blanc
74940 Annecy le Vieux, France
Email: {nicolas.voiron, alexandre.benoit}@univ-smb.fr
patrick.lambert@univ-smb.fr

Andrei Filip and Bogdan Ionescu
LAPI, University Politehnica of Bucharest
061071, Bucharest, Romania
Email: afilip@alpha.imag.pub.ro
bionescu@alpha.imag.pub.ro

***Abstract*—In our data driven world, clustering is of major importance to help end-users and decision makers understanding information structures. Supervised learning techniques rely on ground truth to perform the classification and are usually subject to overtraining issues. On the other hand, unsupervised clustering techniques study the structure of the data without disposing of any training data. Given the difficulty of the task, unsupervised learning tends to provide inferior results to supervised learning. To boost their performance, a compromise is to use learning only for some of the ambiguous classes. In this context, this paper studies the impact of pairwise constraints to unsupervised Spectral Clustering. We introduce a new generalization of constraint propagation which maximizes partitioning quality while reducing annotation costs. Experiments show the efficiency of the proposed scheme.**

***Keywords*—Graph Cut, Spectral Clustering, semi-supervised learning, pairwise constraints, video clustering.**

I. Introduction

Thanks to Internet, the amount of available multimedia information has exploded. One has to rely now on automatic tools to index and classify these huge data in order to provide users with searching and browsing capabilities. In the searching for relevant information, user preferences and suggestions can improve the results by adapting them to each personal's interests. When assessing resembling multimedia content, many approaches have been explored to measure the similarity between images [1] and video [2]. In this paper, we explore these concepts in the context of clustering techniques.

There is currently a lot of literature on clustering [3], [4]. Classic clustering deals with convex data clusters such as the simple k-means algorithm, while more complex approaches, such as mixture-resolving and mode-seeking approaches or artificial neural networks, are able to cope with more difficult cluster representations. One particular category of clustering are the Spectral Clustering Graph Cut techniques [5], that belong to manifold learning. These methods in particular are known to be very effective in dealing with non-convex data clusters. However, standard Spectral Clustering remains unsupervised and cannot benefit from external user knowledge. Recent advances [6], [7] have shown the benefits of introducing pairwise constraints to guide the clustering procedure. Such approaches are close to classic supervised techniques such as Support Vector Machines (SVM), but in contrast, the knowledge is exploited to clustering in a different way.

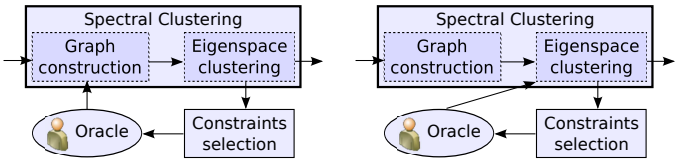
When addressing a complex clustering scenario (e.g., video data), introducing supervision in Spectral Clustering can highly improve clustering performance by solving efficiently the cluster ambiguities. Such constraints, commonly known as "Must Link" and "Cannot Link", indicate if two objects belong or not to the same class. These constraints are generic enough and can be provided via external knowledge, e.g., user input, user studies, etc. Furthermore, similarity annotation is easier compared to absolute class annotation, as we only need to know if two objects belong to a same class. In view of this idea, it is of major interest to optimize the constraints thus to maximize clustering quality while minimizing the costs of user knowledge acquisition. One of the most common strategies [8] consists in using a pairwise constraints automatic propagation approach. However, this propagation strategy is few mentioned in literature and only partially applied.

In this paper, we revisit the concept of pairwise constraint automatic propagation and introduce a new generalization of the bi-partitioning propagation rules for multi-partitioning. Experimentations conducted on two types of datasets, one artificially generated in view of controlling the class separation quality and a publicly available video genre classification dataset [9]; show the efficiency of the proposed approach.

The remainder of the paper is organized as following. Section II presents the current state of the art on Spectral Clustering. Section III-A introduces our contribution. Section IV presents the experimental results while Section V concludes the paper and discusses future work.

II. Semi-supervised Spectral Clustering

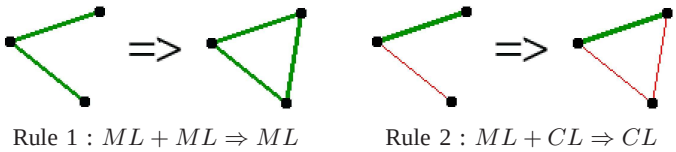
The performance of the data clustering is highly dependent on the properties of the initial data. When dealing with convex data, standard approaches, such as k-means, are able to provide accurate results. However, these approaches cannot deal with data mapped onto complex manifolds that require more elaborated solutions. In the context of similarity graph clustering, state of the art manifold unfolding methods are typically isomap, multidimensional scaling and spectral clustering. These methods generally attempt to identify a lower dimensional space to represent and separate accurately the initial data. In this study, we focus solely on spectral clustering which is able to achieve high accuracy without any assumption on the cluster shape. This method can deal with large datasets when working on sparse similarity graphs, so it is best candidate for approaching large scale video clustering.



(a) Updating initial similarity graph with constraints

(b) Introducing constraints into the clustering problem

Fig. 1: Guiding spectral clustering with the help of the oracle. How to introduce oracle Constraints ?



Rule 1 : $ML + ML \Rightarrow ML$ Rule 2 : $ML + CL \Rightarrow CL$

Fig. 2: Constraint propagation.

However, in its standard form, Spectral Clustering is unsupervised and only relies on the input data. In complex situations, such as video understanding and classification, the semantic gap between extracted features and the expected high-level clustering is significant. Then, the introduction of external knowledge to guide Spectral Clustering becomes of interest. Following this idea, we focus our study on semi-supervised Spectral Clustering by adding sparse pairwise constraints. By accurately choosing some specific constraints in the entire dataset, good clustering results can be expected [10]. This differs from supervised learning which involves the knowledge at training stage being subject to data overtraining.

A. Spectral Clustering

The Spectral Clustering is well described in several tutorials like [5]. It follows three steps.

1) Similarity graph construction: A similarity graph between samples is first constructed. This similarity graph is generally sparse by the usage of a threshold or a k-NN step. Some kernels, like Gaussian one can be also used to expand values.

2) Eigenspace construction: Samples are then projected onto a spectral space where clusters are easier to identify. It is generally a similarity matrix Lapacian eigeinspace.

3) Data clustering in eigenspace: Finally, a standard convex clustering is performed in this space to achieve the final results. State of the art methods generally use a simple k-means. However, other convex clustering methods can also be adopted, e.g., use of Gaussian Mixture Models [6].

B. Pairwise semi-supervision in Spectral Clustering

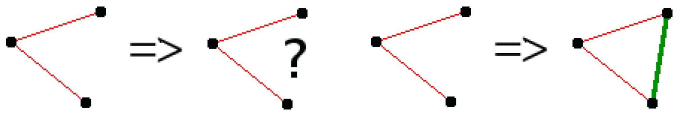
As reported in [6], clustering may introduce semantic ambiguities. This is common in image categorization since the input features are typically low-level descriptors while the targeted clustering is highly semantic. In this context, a solution is the addition of link constraints provided by external knowledge. One approach is in introducing “Must Link” (ML) and “Cannot Link” (CL) constraints between entities. However, the location and number of such constraints should be optimized to enforce clustering quality while keeping a low computational cost.

1) Constraints management: Constraints can be introduced in the first two steps of the spectral clustering (see Section II-A), namely: (i) at *the similarity graph construction* step, as illustrated in figure 1.a, e.g., in [6], inspired by Spectral Learning (SL) [11], authors propose to identify the most ambiguous entities and add into the adjacency matrix, W, their related constraints: ML (1 valued) and CL (0 valued). However, there is no guaranty for the constraints to be followed; or (ii) at *the eigenspace computation step*, as illustrated in figure 1.b. Several approaches were investigated, e.g., “Flexible Constrained Spectral Clustering” (CSP) [12] — introduces the constraints at this step and ends by k-means; “Spectral Clustering with Linear Constraints” (SCLC) [13] — does not use a k-means step but only allows binary clustering; “Constrained Clustering via Spectral Regularization” (CCSR) [14] — adds constraints at an intermediate stage that modifies the eigenspace. A different approach is proposed in [7], “Constrained 1-Spectral Clustering” (COSC), which integrates the constraints into the graph cut problem and is solved in a convex optimization fashion. This directly leads to a two class clustering result without the need of k-means. This method can be extended to multi-class clustering if used in a recursive way. As shown in [7], the COSC method achieves a lower misclassification rate than the previously presented methods in 2 or 10 classes problems. However, performance comparison does not involve constraint propagation but relies on constraints random selection. In a two classes problem with n points, constrained spectral clustering is conducted and ends with a number of pairwise constraints higher that n. However, in such situation, we can expect than n should be the maximum number of constraints needed to obtain the perfect clustering. This motivates us to investigate efficient constraint propagation.

2) Constraints selection: Most of the existing approaches, choose the constraint location randomly. However, [10] showed that the selection of unappropriate constraints may degrade clustering quality while adequate selection can significantly improve it. In general, there are two commonly used metrics employed to identify the relevance of a constraint: the informativeness — measures the amount of the added information, and the consistency — that compares constraints with each others.

The authors in [8] propose an active constraint selection model. It identifies and sorts critical edges on the k-nearest neighbor graph and relies on expert opinion to assess the importance of the ones that link highly connected groups. A similar approach [6] proposes to identify the most ambiguous objects whose links are evaluated by an expert. The originality of this work is in the fact that this object selection tries to focus on the links that have the highest chances to bring significant changes in the Spectral Clustering results.

These methods are in general iterative processes with “a posteriori” constraint selection. This implies an iterative process that involves a clustering step and uses its results at each iteration to select constraints. It thus leads to a high computational cost but proved to be efficient from a partition



(a) Rule 3 in a multi-class case: $CL + CL \Rightarrow \ ?$

(b) Rule 3 in a 2 class case: $CL + CL \Rightarrow ML$

Fig. 3: $CL + CL$ constraint propagation.

quality point of view.

3) Automatic constraints propagation: Once a constraint is introduced, other relations can benefit from this ground truth information and neighborhood ambiguities can be solved without the need of supplementary expert ground truth. This process is illustrated in Figure 2, that shows the transitivity properties of ML+ML and ML+CL combinations of local relations. The propagation is of major interest to lower the cost of expert data and reach a better partitioning quality at a lower computational cost.

III. A CONSTRAINT PROPAGATION FRAMEWORK

A. Constraint propagation generalization

We propose a new way of applying constraints propagation that exploits the combination of two CL constraints. In general, in the multi-class scenario, this combination is uncertain. This case is illustrated in Figure 3.a ($CL + CL \Rightarrow \ ?$). However, as presented in [15], the bi-partition case solves the uncertain case where $CL + CL \Rightarrow ML$, see Figure 3.b. Indeed, in the literature, 2-classes case are often considered. For example, in [7], the COSC method is presented and 3 out of 5 benchmarks are related to bi-partitioning cases. However, in such studies, no propagation is considered.

To our knowledge, in 3 or higher partition cases, rule $CL + CL$ is still reported to be uncertain. However, we can still benefit from this configuration. As shown in Figure 4, on a tetrahedron in a 3 partition case, if 5 CL edges exist, then the last edge is automatically ML. Indeed, with the configuration from Figure 4, objects W and X belong to two different classes; and as Y and Z belong to another class, necessarily Y and Z will belong to the same third class.

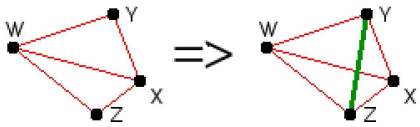


Fig. 4: Example of a 3-class tetrahedron with 5 existing CL constraints, last link is ML.

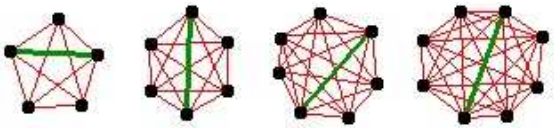


Fig. 5: Example of a n-simplex case with n classes with $\frac{n(n-1)}{2} - 1$ existing CL constraints, last link is ML.

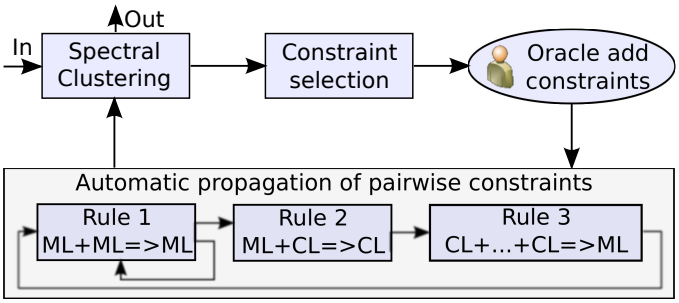


Fig. 6: Overview of the complete automatic pairwise constraint propagation system.

One can generalize this example to the n-simplex of a n partitioning case. If $\left(\frac{n(n-1)}{2} - 1\right)$ CL edges exist, then, the last $\frac{n(n-1)}{2}$th edge is a ML constraint. Demonstration is iterative and analog to the tetrahedron case. Figure 5 illustrates the n-simplex cases with n in range [4,7].

From a computational point of view, high n-order cases are difficult to find and may be rare in sparse similarity matrices but lower orders can be numerous and bring relevant information.

B. Our framework proposal

Figure 6 presents the overall systems that includes spectral clustering and the proposed automatic pairwise constraints propagation system. Once a spectral clustering iteration is applied, oracle assesses a set of constraints, and propagation is applied using the following procedure:

- Rule 1 is first used : MLs are introduced and combined with the existing ones to generate new ones. This process is applied iteratively up to idempotence.
- Rule 2 is applied next and generates several Cl constraints in a single pass.
- Rule 3 is applied to generate new ML constraints. If at least one is generated, the entire process is restarted.

Propagation limit is the fully connected graph. However, in large dataset with very few annotations, propagation stops earlier.

IV. EXPERIMENTAL RESULTS

Experimental validation is conducted on two types of datasets: an artificially generated dataset that allows class separation to be controlled[1] and the real-world Blip10000 [9] video genre classification dataset. Both datasets are used in 2-class or multi-class partitioning configurations. In what concerns the Spectral Clustering approaches, we experimented with 2 state of the art approaches: the Active Clustering method [6] which manages constraints in the similarity matrix construction step and does not ensure constraints to be followed; Constrained 1-Spectral Clustering (COSC) [7] which handles constraints in the eigenspace construction stage and tries to fulfill all of them.

[1]data will be available for the final submission.

To assess performance, we use the standard adjusted Rand index [16] which measures the similarity between two data clustering. Its main advantage is to take its values between -1 and 1 where the value 1 means that the 2 partitions are equals and where mainly the 0 value indicates than the 2 partitions are untied. We compute Rand index between ground truth clusting and the obtain clustering.

We adopted the following validation procedure [6]: a first clustering is applied without any constraint. At each next iteration, some new constraints are chosen randomly and involved in the process and managed by each Spectral Clustering method. To be fair and not algorithm dependant, constraints are chosen randomely from the ground truth. Constraints are propagated recursively ensuring that all possible propagations have been performed at each iteration. Performance of the results is then compared to ground truth clusters with the Rand index.

A. 2-class experiments

In this experiments, no normalization is applied and computed similarities are built from the symmetric 5 nearest neighbors. A Gaussian weighting is employed.

We have generated five sets of data consisting of 100 objects spread over 2 circular classes (see right panels in Figure 7) with the following properties:

- *1st dataset*: has randomly placed objects;
- *2nd dataset*: is partially mixed and has classes with a wide radius variance while their support overlaps;
- *3rd dataset*: has classes with wide variance radius and non overlapping but contiguous areas;
- *4th dataset*: similar to third one but areas remain separated;
- *5th dataset*: has strongly separated classes.

Figure 7 presents the results of each clustering method for each of the considered data sets while using either no constraint propagation (black curves), the two first propagation rules (blue curves) or all of the three rules (red curves). Each curve is the average of 20 different randomly generated runs. One can observe that Active Clustering benefits the most from constraint propagation. COSC method performs in general better and only takes advantage of the constraint propagation when dealing with complex situations, e.g., 1st and 2nd datasets.

From a general point of view, the use of all the three rules ensures to reach perfect partitioning in less than 100 supervised pairs, i.e., 2% of all the possible relations. In difficult cases, e.g., 1st and 2nd datasets, on average, COSC is able to converge two times faster when using constraint propagation compared to the original solution. Active Clustering converges slower than COSC but its speed factor provided by the use of constraint propagation is higher. As a reference, in an easy case such as for the 5th dataset, both clustering methods directly converge to maximum partition quality without the need of constraints.

One can also observe on the 2-complex dataset examples that partition quality starts to increase only after a significant

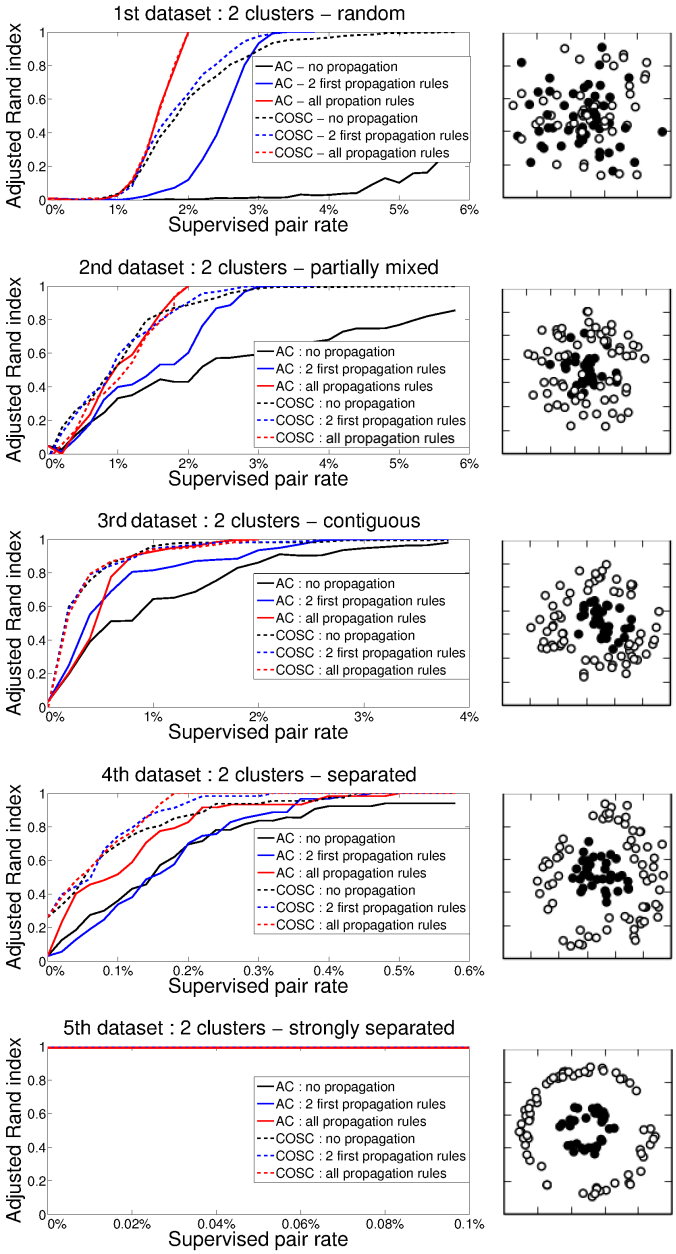


Fig. 7: Partition quality as a function of the supervised pair rate for Active Learning (continuous lines) and COSC (dashed lines) using no propagation and the 2 first or 3 propagation rules. Results are reported on 5 reference datasets having an increasing class separation configuration.

addition of constraints. This is related to the effectiveness of constraint propagation as a function of supervised pairs as illustrated in figure 8. When constraints are spread all over the dataset, for example when selecting constraints randomly, too few pairs are connected which limits propagation. Propagation increases much faster as long as constraints connections appear.

Regarding the impact of the third rule that we propose, the first graph of the figure 8 shows its effect onto the overall propagation system in a two class problem. The two solid lines

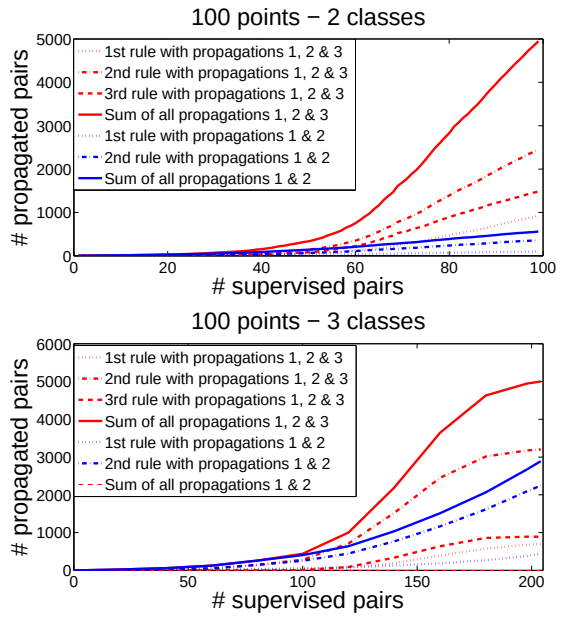


Fig. 8: Number of propagated constraints as a function of supervised pairs for a 100 points 2 or 3 class dataset.

show the total number of added constraints with respect to a set of used rules. Blue solid curve only relies on the use of the two first rules while the red one shows results obtained with all the three rules. This latter case shows a significant increase in constraint addition. More into details, those added constraints do not only rely on the third rule : red solid line is much higher than the sum of the blue solid curve added to the red dotted line of rule 3. Indeed, the total number of added constraints benefits from a domino effect coming from a synergy of all the rules, and the third one has a dramatic effect. The second graph of the figure 8 shows a similar tendency in the case of a 3 class problem. In such case, the number of constraints added by the third rule is already significant but its effect on the total number of added constraints is even higher. In this last example, once the oracle assessed 140 pairs, using only the two first rules allowed 1035 constraints to be added. Using also the third rule, 338 constraints are directly added, but 2191 are globally added. Finally, the third rule allows a constraints gain of 110% compared to the use on only the two first rules. Also, coming back to the oracle work, its contribution is amplified by a factor of 15.6 by the global framework.

B. Multi-class experiments

We first present the results obtained on a 100 point controlled dataset which is divided into 3 equivalent classes. Each points of each class are randomly placed in a unity circle. The second dataset is composed of 100 points corresponding to real video sequences from the Blip10000 [9] dataset. Each video is described via the *standard audio features* proposed in [17]. In this dataset, videos are classified in 3 equal sized categories: “Health”, “Documentary” and “Literature”. A third dataset is proposed to show the data scalability potential of these methods. It is similar to the second one but is much larger and comprises 5,197 videos from the Blip10000 dataset that are distributed in 26 video genre categories. Clusters have

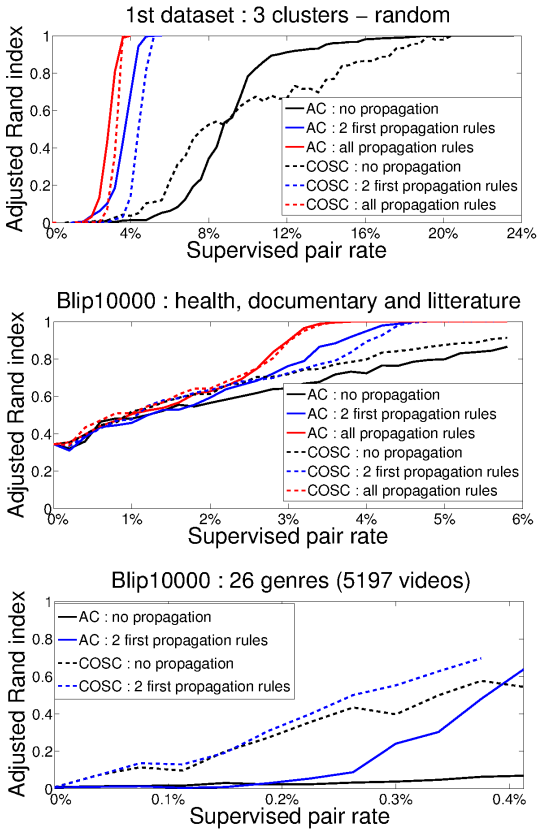


Fig. 9: Partition quality as a function of pairwise constraints addition. Continuous lines for Active Clustering method and dashed lines for COSC method. Red lines show the 3 rules propagation results, blue shows the 2 first propagation rules results. Black lines are original method results.

a variable number of items.

Results are presented in Figure 9. One can observe that in this new use case, automatic propagation brings significant convergence speed gain. More specifically, the two first datasets show the advantage of the third propagation rule proposed in Section III-A which allows perfect partitioning with 20% less constraint addition costs compared to the limited use of the two first propagation rules. In this experiment, COSC achieves lower performance than Active Clustering. This can be explained by the fact that COSC hierarchically clusters data in 2-classes problems which is less adapted to this 3-classes context.

When considering larger scale datasets with large number of clusters, COSC outperforms Active Clustering. Both methods benefit from the proposed constraint propagation.

Results shows that the 2 first propagation rules applied on 50,000 links, i.e., 0.37% of all possible links of the graph, allow partition quality to be improved by 21% for COSC and by 650% for Active Clustering. However, in such cases, one should discuss about the computational costs of the propagation technique. Indeed, the two first rules can be applied efficiently thanks to matrix operation vectorization. But the third rule consists here in the analysis of all the 26-simplex analysis. For now, such analysis is not optimized and becomes

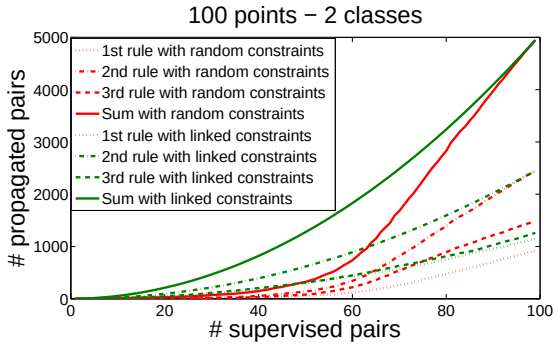


Fig. 10: Number of propagated constraints as a function of supervised pairs for a 100 points 2 class dataset. Both for total random and random linked selection constraints.

too much time and memory consuming. As a consequence, only the two first rules have been applied. Further work should try to improve on this aspect to support large scale analysis.

C. Amplifying propagation effect with a constraint selection strategy

From the beginning, object pairs were selected randomly from the unsupervised pairs set. However, as we have seen in section II-B2, many constraint selection strategies exist and they have chances to amplify even more the effect of the proposed propagation method. A simple yet efficient way to do it is to restrict random pair selection to the subset of pairs having only one object connected to an already supervised pair. Figure 10 compares this "random linked selection" strategy to the "totally random" selection on the same 2 class dataset as the one used in figure 8. On green curves we can see that the random linked selection strategy boosts propagation for all the rules from the beginning. As an illustration, once 20 pairs have been assessed by the oracle, the totally random constraint strategy allowed 30 constraints to be added while random linked constraint strategy added 210. Constraints gain amplification factor is here 7. This strategy is interesting when oracle cost is high and forces annotation process to stop early. However, in the end, both selection methods converge to the same value.

V. Conclusions

This paper presented a generalization of constraint propagation used in the case of similarity graph clustering optimization. We experimented with a controlled dataset and real-world video genre classification problem and showed the repeatable benefit of the use of the proposed constraints propagation techniques. Two different state of the art Spectral Clustering techniques that manage constraints in different ways benefit from it and converge faster to higher partition quality.

The main contribution of this paper is in the generalization of constraint propagation techniques. We do recommend its usage since it reduces the cost of constraints addition to enhance clustering quality. In the worst case, performances would be the ones of the original clustering methods.

Further work will mainly address the optimizing of the constraint propagation from a computational point of view. Also, the constraints selection topic becomes of major interest in maximizing methods efficiency. An other perspective is to compare such approach to state of the art supervised clustering techniques in challenges such as MediaEval.

References

[1] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Comput. Surv.*, vol. 40, no. 2, pp. 5:1–5:60, May 2008. [Online]. Available: http://doi.acm.org/10.1145/1348246.1348248

[2] N. Voiron, A. Benoit, and P. Lambert, "Automatic difference measure between movies using dissimilarity measure fusion and rank correlation coefficients," in *Content-Based Multimedia Indexing (CBMI), 10th International Workshop*, 2012.

[3] J. Camargo and F. González, "Visualization, summarization and exploration of large collections of images: State of the art," in *Latin-American Conference On Networked and Electronic Media. LACNEM*, 2009.

[4] I. Witten, E. Frank, and M. Hall, "Data mining: Practical machine learning tools and techniques," *Morgan Kaufmann Publishers*, 2011.

[5] U. von Luxburg, "A tutorial on spectral clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, Dec. 2007. [Online]. Available: http://dx.doi.org/10.1007/s11222-007-9033-z

[6] C. Xiong, D. M. Johnson, and J. J. Corso, "Active clustering with model-based uncertainty reduction," *CoRR*, vol. abs/1402.1783, 2014.

[7] S. S. Rangapuram and M. Hein, "Constrained 1-spectral clustering," in *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics, AISTATS 2012, La Palma, Canary Islands, April 21-23, 2012*, 2012, pp. 1143–1151. [Online]. Available: http://jmlr.csail.mit.edu/proceedings/papers/v22/sundar12.html

[8] V. Vu, N. Labroche, and B. Bouchon-Meunier, "Improving constrained clustering with active query selection," *Pattern Recognition*, vol. 45, no. 4, pp. 1749–1758, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.patcog.2011.10.016

[9] S. Schmiedeke, P. Xu, I. Ferrané, M. Eskevich, C. Kofler, M. Larson, Y. Estève, L. Lamel, G. Jones, and T. Sikora, "Blip10000: A social video dataset containing spug content for tagging and retrieval," *ACM Multimedia Systems Conference*, 2013.

[10] I. Davidson, K. L. Wagstaff, and S. Basu, "Measuring constraint-set utility for partitional clustering algorithms," in *In: Proceedings of the Tenth European Conference on Principles and Practice of Knowledge Discovery in Databases*. Springer, 2006, pp. 115–126.

[11] S. D. Kamvar, D. Klein, and C. D. Manning, "Spectral learning," in *In IJCAI*, 2003, pp. 561–566.

[12] X. Wang and I. Davidson, "Flexible constrained spectral clustering," in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, July 25-28, 2010*, 2010, pp. 563–572. [Online]. Available: http://doi.acm.org/10.1145/1835804.1835877

[13] L. Xu, W. Li, and D. Schuurmans, "Fast normalized cut with linear constraints," *2013 IEEE Conference on Computer Vision and Pattern Recognition*, vol. 0, pp. 2866–2873, 2009.

[14] Z. Li, J. Liu, and X. Tang, "Constrained clustering via spectral regularization," in *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009), 20-25 June 2009, Miami, Florida, USA*, 2009, pp. 421–428. [Online]. Available: http://dx.doi.org/10.1109/CVPRW.2009.5206852

[15] P. K. Mallapragada, R. Jin, and A. K. Jain, "Active query selection for semi-supervised clustering," in *19th International Conference on Pattern Recognition (ICPR 2008), December 8-11, 2008, Tampa, Florida, USA*, 2008, pp. 1–4. [Online]. Available: http://dx.doi.org/10.1109/ICPR.2008.4761792

[16] L. Hubert and P. Arabie, "Comparing partitions," *Journal of classification*, vol. 2, no. 1, pp. 193–218, 1985. [Online]. Available: http://scholar.google.de/scholar.bib?q=info:IkrWWF2JxwoJ:scholar.google.com/&output=citation&hl=de&ct=citation&cd=0

[17] I. Mironica, B. Ionescu, P. Knees, and P. Lambert, "An in-depth evaluation of multimodal video genre categorization," in *IEEE International Workshop on Content-Based Multimedia Indexing*, 2013.